

明 細 書

ロボットの駆動方法

技術分野

- [0001] 本発明は、ロボットの駆動方法に関するものであり、特にロボットを動作させるためにそのロボットに対応する動作物またはその擬似物に所望の動作を行わせ、その動作物または擬似物の画像データと同じ動作をロボットに行わせるのに適したロボットの駆動方法に関するものである。

背景技術

- [0002] 例えば、人の動作を見まねで再現できるロボットに関する技術としては、操作者がセンサを付けて歩行したときや、操作者が複数のセンサを備えたデータグローブを手 に嵌めて手を動かしたときに、それらのセンサから出力される情報に基いてロボットを駆動している。非特許文献1及び非特許文献2には、この種の技術が開示されている。

非特許文献1: Y. Kuniyoshi著の“A robotic approach to the foundation of communication and symbols”と題する論文: Proc. of 12th World Congress of Applied Linguistics (AILA'99), 94頁, 1999年発行

非特許文献2: T. Inamura, Y. Nakamura, H. Ezaki及びI. Toshima著の“Imitation and primitive symbol acquisition of humanoids by the integrated mimesis loop”と題する論文: Proc. of Intl. Conf. on Robotics and Automation (ICRA 2001), 4208頁-4213頁, 2001年発行

発明の開示

発明が解決しようとする課題

- [0003] しかしながら、上記従来の技術では、多くの場合、不良設定性(センサの出力または認識によって正確に判別できない動き)がほとんど存在せず、解が一意に決まり、比較的容易にロボットを駆動することができる。一方、不良設定性が存在するような問題に対しては、常にロボットに対応する動作物またはその擬似物にセンサを装着しておく必要があり、汎用性に欠ける問題がある。

- [0004] 本発明の目的は、実際にロボットを駆動する際に、ロボットに対応する動作物またはその擬似物にセンサを装着しておく必要のないロボットの駆動方法を提供することにある。
- [0005] 本発明の他の目的は、動作画像データを入力データとしてロボットを駆動することを可能にするロボットの駆動方法を提供することにある。
- [0006] 本発明の他の目的は、動作物またはその模擬物の形状が複雑であったり、画像上の自己遮蔽が頻繁に起こる場合でも、安定して高速にロボットを駆動することを可能にするロボットの駆動方法を提供することにある。
- [0007] 本発明の更に他の目的は、コンピュータグラフィック画像データを利用してロボットを駆動することが可能なロボットの駆動方法を提供することにある。
- [0008] 本発明の別の目的は、入力される動作画像データの特定精度を高めることができるロボットの駆動方法を提供することにある。
- [0009] 本発明の他の目的は、入力された動作画像データと事前に記憶した多数の画像データとの照合を迅速に行うことができるロボットの駆動方法を提供することにある。
- [0010] 本発明の更に他の目的は、見まね学習により不良設定性を解消してロボットを駆動する方法を提供することにある。

課題を解決するための手段

- [0011] 本発明のロボットの駆動方法は、次のステップ従って、動作指令によりロボットを駆動する。まずロボットに対応する動作物またはその擬似物に所定の動作と同じ動作をさせたときの動作物またはその擬似物の複数の画像データと、これら複数の画像データに対応した動作指令を得るために事前に収集された事前収集動作指令とを対応付けて画像対応動作指令記憶手段に記憶する。事前収集動作指令を事前に収集するための方法は任意である。次に、ロボットを動作させるために、動作物またはその擬似物に所望の動作を行わせ、その際に動作物またはその擬似物の動作画像データを時系列でロボット動作画像データとして取得する。そしてロボット動作画像データに含まれる動作画像データに対応する画像データを、画像対応動作指令記憶手段に記憶している複数の画像データから時系列で特定し、特定した画像データに対応する事前収集動作指令を動作指令としてロボットに与える。なお画像対応

動作指令記憶手段に記憶する画像データは、撮影した画像に基づいて処理を行って得たその画像に関する特徴を示す特徴量のデータを伴うものまたは特徴量のデータそのものであってもよいのは勿論である。そしてロボット動作画像データに含まれる動作画像データに対応する画像データの特定方法は、任意である。画像データが特徴量のデータであれば、特徴量の類似度によって、動作画像データに対応する画像データの特定を行うことができる。

- [0012] 本発明の方法により実際にロボットを動かすと、ロボットに対応する動作物またはその擬似物の画像を得るだけで、センサを使用することなく、その動作物または擬似物と同じ動作を、ロボットに行わせることができる。したがって本発明によれば、画像データを入力データとしてロボットを簡単に駆動することが可能になる。
- [0013] なおロボット動作画像データに含まれる画像データと画像対応動作指令記憶手段に記憶されている複数の画像データとの対応を、両者の類似度に基づいて判断することができる。その場合には、画像対応動作指令記憶手段に記憶している複数の画像データから動作画像データに対応する画像データを特定する際に、動作画像データの特徴量を基準にして照合用の複数の画像データを選択し、照合用の複数の画像データと動作画像データとの類似度に基づいて動作画像データに対応する画像を特定することが好ましい。
- [0014] 本発明のより具体的な方法では、以下の第1乃至第5のステップを実行することによりロボットを駆動する。第1のステップでは、ロボットに対応する動作物と、この動作物に設けられて動作物の動きを検出する複数のセンサと、これら複数のセンサの出力に基づいて動作指令を発生する動作指令発生装置とを用いる。そして動作物に所定の動作を行わせたときの複数のセンサの出力に基づいて動作指令発生装置が発生する動作指令を事前収集動作指令として記憶する。ここで動作物とは、典型的には人間全体または人間の手や足等の動作する部位等であるが、動作物は人間以外の動物等でもよい。
- [0015] 第2のステップでは、動作物に第1のステップにおける所定の動作を行わせたときの動作物の複数の画像データを取得するか、または動作物またはその擬似物に前記所定の動作と同じ動作をさせたときの動作物またはその擬似物の複数の画像データ

を時系列で取得する。ここで「動作物またはその擬似物」における「動作物」とは第1のステップの動作物と同じものでもよいが、同様の形状及び構造を有し(言換えると、同じ場所に同じ関節があり)且つ同じ動作を行い得るものであれば別の動作物であってもよい。また「動作物またはその擬似物」における「その擬似物」とは、その動作物をコンピュータグラフィック技術等の擬似物生成技術を用いて作成したものや、その動作物の人形等である。この場合に得る画像データは、疑似画像データである。特にコンピュータグラフィック技術を用いる場合の疑似物の画像データは、コンピュータグラフィック画像データである。複数の画像データを時系列で取得する方法または手法は、任意であり、単眼カメラ、両眼カメラ、あるいはマルチカメラのいずれにより撮影したものであってもよい。

[0016] 動作物が人間の手である場合、第2のステップで、複数の画像データに、人間の手に現れる個人差を考慮した画像データを含めておくと、後述する第5のステップにおける画像データの特定精度を高めることができる。ここで個人差が生じる相違要素には、拇指以外の4指の基節位置の曲率、拇指基節の張り出し具合、隣り合う二本の指の開き具合、開いた二本の指の基準角度の違い、拇指の反り具合等がある。このような個人差は、コンピュータグラフィック技術等の擬似物生成技術を用いて画像データ(疑似画像データ)を作成する場合には、前述の相違要素をパラメータとして既に作成した基本となる画像データを変更することにより(骨の長さや関節の可動角度といったような、前述の相違要素を生み出しているパラメータを変更することにより)簡単に得ることができる。したがって個人差を考慮した画像データの作成は、難しいものではなく、簡単である。

[0017] また第2のステップで作成する複数の画像データには、複数の画像データの解像度を変えて作成した複数の解像度変更画像データを含めてもよい。これは画像データの解像度を変えたほうが、画像の特徴が現れやすくなるものもあるからである。画像データの解像度は、基本となる画像データがあれば、簡単に変更することは可能である。

[0018] なお第2のステップでは、複数のセンサを備えた動作物の表面を、複数のセンサを含めて被覆物によって覆い、第1のステップと同時に動作物の複数の画像データを

取得するようにしてもよい。被覆物によって動作物を覆うと、動作物の表面の不要な情報を物理的に取り去ることができる。したがって被覆物は、不要な情報をできるだけ排除できる構成を有しているのが好ましい。そのためできるだけ無地(一色で無模様:不要な凹凸がないもの)であることが好ましい。このような被覆物で覆った動作物の画像データからは、その画像データに含まれる特徴(特に輪郭についての特徴)を短い時間で、しかも高精度に取得することができる。

[0019] 第3のステップでは、第2のステップで得た複数の画像データに含まれる各画像データと事前収集動作指令とを対応付けて第2のステップで得た複数の画像データと事前収集動作指令とを画像対応動作指令記憶手段に記憶する。なお第3のステップでは、実際に撮影した画像だけでなく、第2のステップで取得した前の画像データと後の画像データとの間の未作成(未撮影)の画像データをコンピュータグラフィック技術により作成したものも記憶の対象とすることができる。この場合には、未作成の画像データに対応する事前収集動作指令を、前の画像データに対応する事前収集動作指令と後の画像データに対応する事前収集動作指令とに基づいて推定により作成し、これを未作成の画像データと対応付けて記憶することができる。このようにすれば、少ない実際の画像データを用いて、より多くの画像データと事前収集動作指令との対応データを得ることができ、基礎データ(画像データと事前収集動作指令とを対応付けたデータ)の収集が大幅に簡単になる。

[0020] 第4のステップでは、ロボットを動作させるために、動作物またはその擬似物に所望の動作を行わせ、その際に動作物またはその擬似物の動作画像データを時系列でロボット動作用画像データとして取得する。この場合に、動作物にセンサを装着する必要はない。そして第5のステップでは、ロボット動作用画像データに含まれる動作画像データに対応する画像データを画像対応動作指令記憶手段に記憶している複数の画像データ(基礎データ)から時系列で特定する。なおロボット動作用画像データに含まれる動作画像データと画像対応動作指令記憶手段に記憶されている複数の画像データとの対応は、両者の類似度に基づいて判断することが好ましい。類似度の求め方は任意であり、特定の方法に限定されるものではない。そして第5のステップでは、特定した画像データに対応する事前収集動作指令を動作指令としてロボ

ットに与える。

[0021] なお人間型ロボットにおいて不良設定性が多いのは、手である。したがってロボットハンドの駆動に本発明を適用すると、従来よりも簡単にロボットハンドを駆動することができる。なおこの場合、動作物としては人間の手またはその擬似物が用いられることになる。そして第1のステップでは、人間の手に着用するデータグローブを用いる。このデータグローブは、グローブ本体にロボットの手の動作部に対応した人間の手の動作部の動きを検出する位置に複数のセンサが装着された構造を有している。

[0022] 基礎データとなる画像データが多くなるほど(画像データベースのデータ量が多くなるほど)、基礎データとなる画像データの中から、入力である動作画像データに対応した画像データを特定するには時間がかかるようになる。そこでこのような問題を解消するためには、第3のステップで、画像データと事前収集動作指令とを対応付けて画像対応動作指令記憶手段に記憶する際に、次のようにすることができる。まず複数の画像データのそれぞれの特徴量を個々に演算する(特徴量ステップ)。次に複数の画像データのそれぞれの特徴量を主成分分析して、複数の画像データのそれぞれについて主成分ごとの主成分得点を演算し、且つ累積寄与率に基づいて第1主成分から第k主成分までの主成分の数を決定する(主成分決定ステップ)。そして第1主成分から第k主成分までの各主成分ごとに、複数の画像データを主成分得点の大きさを基準にして並べ替えて得たk種類の画像データソース[画像情報の表(テーブル)]を作成して記憶する(記憶ステップ)。すなわち例えば、15,000枚分の画像データがあるとした場合に、10種類の主成分を決定したとすると、15,000枚分の画像データを10種類の主成分ごとに、その主成分得点の大きさを基準にして並べ替えた10種類の画像データソース($15,000 \times 10 = 150,000$ 枚分画像データに相当)として記憶する。画像データソース内の並べ変えた各画像データは、それぞれ事前収集動作指令と対応付けられて記憶されている。このようにした場合には、第5ステップで、入力された動作画像データについても主成分得点を求め、この主成分得点を基準にしてk種類の画像データソースから照合用の複数の画像データをそれぞれ選出する。動作画像データの解像度を変えた複数種類の解像度の異なる動作画像データを作り、この複数種類の解像度の異なる動作画像データについて主成分得点を求

め、この主成分得点を基準にして k 種類の画像データソースから照合用の複数の画像データをそれぞれ選出するようにしてもよい。この照合用の複数の画像データは、各画像データソースから各主成分ごとの主成分得点を基準にして所定枚数分ずつ選出される。このようにすると記憶されているすべての画像データと動作画像データとを照合する場合と比べて、照合対象となる画像データの数を少なくすることができるので、照合時間が短くてすむ利点が得られる。そして主成分得点を基準にして動作画像データと照合用の複数の画像データの類似度を判定し、前記類似度に基づいて動作画像データに対応する画像データを特定するようにしてもよい。このように解像度の異なる動作画像データを作成してそれらの因子負荷量も利用して類似度の判定を行うと、画像の特定精度を上げることができる。

- [0023] なお第1のステップは事前に行う作業であるため、第1のステップを切り離して、ロボットを駆動する装置に用いられる駆動方法においては、第2乃至第5のステップだけを用いてロボットを駆動することができるのは当然である。

図面の簡単な説明

- [0024] [図1]発明者等が開発した本実施の形態で制御の対象とする公知の多指ロボットハンド1の平面図である。
- [図2]図1のロボットハンドの内部に配置されたアクチュエータの配置構成を示す図である。
- [図3]サイン波入力に対する追従性能を示す図である。
- [図4]サインFM変調波入力に対する追従性能を示す図である。
- [図5]遠隔指令装置として使用するデータグローブの一例の概略構成を示す図である。
- [図6]多指ロボットハンドの制御装置及び本発明の駆動方法を実施する装置の一例の概略構成を示す図である。
- [図7]図6の対応関係決定手段の主要部をコンピュータを用いて実現する場合に用いるソフトウェアのアルゴリズムを示すフローチャートである。
- [図8]第1乃至第3のステップをコンピュータを用いて実現する場合に用いるソフトウェアのアルゴリズムを示す図である。

[図9]第1乃至第3のステップを説明するために用いる図である。

[図10]第4及び第5のステップをコンピュータを用いて実現する場合に用いるソフトウェアのアルゴリズムを示す図である。

[図11]高次局所自己相関特徴のための局所パターンの例を示す図である。

[図12](A)乃至(C)は、人の手に現れる個人差の相違要素を示している。

[図13]コンピュータグラフィック編集ソフトのパラメータを変えて、基節位置の曲率と拇指基節の張り出し具合を変えた場合の画像データの相違を示す図である。

[図14]解像度を変えた3種類の画像データを示す図である。

[図15](A)及び(B)は、 320×240 画素から構成された1画面を 4×4 画面に分割した状態と、 8×8 画面に分割した状態を示す図である。

[図16]画像データを並べ替えたイメージを示す図である。

[図17]画像データのデータベース作成の手順を示すフローチャートである。

[図18]具体的な照合手順の流れを示すフローチャートである。

[図19]本発明の方法の別の実施の形態を説明するために用いる概念図である。

[図20]データベースを作成するためのアルゴリズムを示すフローチャートである。

[図21]図20のアルゴリズムの詳細を示すフローチャートである。

[図22]輪郭抽出した画像データを示す図である。

[図23]データベース作成に使用する高次局所自己相関特徴のための局所パターンの例を示す図である。

[図24]特徴量の抽出を説明するために用いる図である。

[図25]動作画像データに対応する画像データをデータベースに記憶した画像データから選ぶ際のアルゴリズムを示すフローチャートである。

[図26]図25のアルゴリズムの詳細を示すフローチャートである。

[図27]クラスタリングを説明するために用いる図である。

発明を実施するための最良の形態

[0025] 以下図面を参照して本発明の方法の実施の形態について説明する。

[0026] なお実施の形態について説明する前に、本実施の形態で用いる技術の基礎について説明する。なお実施の形態について説明する前に、本実施の形態で用いる技

術の基礎について説明する。両眼あるいは単眼カメラによるヒト手指の3次元形状を推定する場合には、手指の形状が複雑で、しかも自己遮蔽も多く、形状の認識における不良設定性が存在する。そのため、赤外線レンジセンサや複数台のコンピュータを用いた並列計算装置といった周辺機器類を使用しない限り、高速で且つ高精度に推定することはできない(参考文献:E. Ueda, Y. Matsumoto, M. Imai and T. Ogasawara: “Hand pose estimation using multi-viewpoint silhouette images,” Proc. 2001 IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS’2001), pp.1989-1996, 2001。M. H. Jeong, Y. Kuno, N. Shimada and Y. Shirai: “Recognition of shape-changing hand gestures,” IEICE Transactions Division D, E85-D, 10, pp.1678-1687, 2002。N. Shimada, K. Kimura, Y. Kuno and Y. Shirai: “3-D hand posture estimation by indexing monocular silhouette images,” Proc. 6th Workshop on Frontier of Computer Vision, pp.150-155, 2000)。一方、ヒト型のロボット等の上肢制御には、少なくとも軌道選択、関節角度の組み合わせ、筋張力の組み合わせなどの不良設定問題があることが知られている(Y. Uno, K. Kawato及びR. Suzuki 著の“Formation and control of optimal trajectory in human multijoint arm movement—minimum torque—change model”と題する論文:Biological Cybernetics, 61, pp. 89-101, 1989.)。ところが、両方の不良設定性を、人は巧く解決して、複雑精緻な随意運動を実現している。経験則から言えば、自分自身の身体での運動経験、すなわち身体性(R. Pfeifer, “Dynamics, morphology, and materials in the emergence of cognition,” In Advances in Artificial Intelligence, ed. W. Burgard, A. B. Cremers, and T. Christaller, Proc. KI-99, 23th Annual German Conference on Artificial Intelligence, 1999.)の存在が、認識における不良設定問題の重要な解決になっているということもできる。

- [0027] そこで、発明者らは、認識と制御の両方における不良設定問題を同時に解決できる非常に有効な、場合によっては唯一の方法が、見まね学習であると考えて、人と同じような精度と処理速度で手指動作の再現が可能なロボットハンド・システムの開発を行った。

[0028] 以下、このシステムに基づいて本発明のロボットの駆動方法の実施の形態の幾つかの例を説明する。

[0029] まず本実施の形態で用いる多指ロボットハンド(ロボット)の機構について説明する。図1は、発明者等が開発した本実施の形態で制御の対象とする公知の多指ロボットハンド1の平面図であり、図2は図1のロボットハンドの内部に配置されたアクチュエータ2(エンコーダ内蔵型超小形モータ)の配置構成を示す図である。なおこの多指ロボットハンドの詳細は、特開2003-117873号公報に示されている。この多指ロボットハンド1の各関節の可動範囲は、親指の内外旋が 120° であり、内外転が 60° である。そして各5本指の屈曲、伸展は第2関節に連動して行われ、第2関節の可動範囲は 112° であり、第1関節は第2関節の可動範囲の $7/10$ 倍、第3関節は第2関節の可動範囲の $5/7$ 倍となっている。指の開閉は薬指に連動して、人指し指、小指が動き、中指は動かない。また薬指の開閉可動範囲は 15° であり、人指し指は 13° であり、小指は 31° である。これらの構成によって、拇指と小指の接触が可能になり人の手指の可動域を十分に実現できている。

[0030] 各関節のアクチュエータ2には、DCマイクロモータを使用している。各指の第2関節、拇指の内外旋、内外転部位と拇指以外の4指間の開閉(アブダクション)用に掌の内部にそれぞれアクチュエータ2が設置されている。各関節の制御は、図示しない制御用計算機からの指令を受けて動作し、内部にD/Aコンバータや、小型サーボモータ専用ドライバー(マイクロサーボ)が配置された駆動装置3を介して、各アクチュエータ2のモータにかかる電圧を変化させることで行う。各アクチュエータ2のモータには、インクリメンタルシャフトエンコーダが内蔵されている。このエンコーダの出力は、軸速度、回転方向の検出の他に、位置制御に使用できる。各関節の角度情報は、ギヤを介して分解能を挙げたエンコーダを介して図示しない制御用計算機にフィードバックされる機構を持っている。エンコーダ自体の分解能は $16[\text{Pulse}/\text{Revolution}]$ である。指の開閉部のエンコーダには減速比 $1/400$ のギヤが設けられており、換算分解能は $6400[\text{Pulse}/\text{Revolution}]$ となる。その他のエンコーダには減速比 $1/50$ のギヤを設けており、換算分解能は $800[\text{Pulse}/\text{Revolution}]$ となる。

[0031] PID制御等を用いて、サイン波、FM変調波の時系列角度変化指令を入力したとき

の追従結果を図3及び図4に示す。図3はサイン波入力に対する追従性能を示し、図4はサインFM変調波入力に対する追従性能を示す。これらの結果から、ゆっくりとした動作から比較的早い動作、変調波に対しても良好な追従特性が得られることが分かる。

[0032] 図5は、遠隔指令装置として使用するデータグローブ4の一例の概略構成を示す図である。このデータグローブ4には、手袋形の本体の関節部分を中心にして約16箇所、位置に分散して位置センサが備えられた構成を有している。具体的には、Virtual technologies社製のcyberglove(商標)をデータグローブとして用いている。このデータグローブのセンサ5は、拇指を除く各指のMP(中手指節関節)、PIP(近位指節間関節)に一つずつ、拇指にはMP、PIP、DIP(遠位節間関節)についており、また各指の間にあるアブダクションを計測するセンサと掌の中央のセンサとで計16ヶ所に設置されている各センサ5はそれぞれ歪曲(ひずみ)度を数値化して30〜40Hzで出力する仕様を有している。掌の中央に設置されているセンサ5は掌全体の歪曲度を計測するものであり、対する人型ロボットハンドにはその機構は備わっていないため、この実施の形態では、この中央のセンサ5は使用しない。

[0033] 人の手指動作を操作者が手に装着したデータグローブによりデータ化し、これを計算機に取り込んで多指型ロボットハンドの制御指令値に変換すれば、データグローブを装着した操作者の手の動作と同様の動作をロボットハンドで実現させることができる。

[0034] データグローブによるロボットハンドの制御法において、一つの簡単な方法としては、サイバークロブから出力されるデータを3次元座標位置に変換しロボットハンドの関節と対になる部分の角度情報を時系列で取り出し、それをロボットハンドに実行させて、ロボットハンドを動作させるという手段がある。しかし、この方法を用いるとデータグローブの各センサから出力される16個の値がそれぞれ各関節のX座標、Y座標、Z座標、ピッチ角、ヨー角、ロール角に変換され、一単位時刻あたり100ものデータが生成されることとなり、これからロボットハンドの動作に必要な8関節分のデータを抽出して使用することは計算効率や時間効率を考える面からも最適ではない。

[0035] 人の手指は拇指を除く四指に至って、他の四指の影響を少なからず受ける。その

特徴を定量的に表すこともこの制御方法及び装置において重要であると考え。そこでデータグローブで得られるデータからロボットハンド制御用の指令値への変換式に本発明の実施の形態では、重回帰式を用いる。

- [0036] ロボットハンドの各関節に対し重回帰式、すなわちすべてのデータグローブからの出力の重み付き線形和によりそれぞれのロボットハンドの関節を記述するためには、まず重回帰分析を行い偏回帰係数を求める必要がある。そこで重回帰分析用のデータを取得するためにデータグローブ装着者は予め定めた時系列関節角度パターンに従って関節角度指令値を駆動装置3に入力して、ロボットハンドを動作させる。なお駆動装置3に指令を与える制御装置は、ロボットハンド1にあらかじめ重回帰分析に有効な動作を行わせる予め定めた時系列関節角度パターンの関節角度指令値を記憶している。そしてこのロボットハンドの動作に従い即ち倣って、データグローブを装着した操作者には、データグローブを装着した手において、指の屈伸、アブダクション等の動作を事前に行ってもらふ。本発明では、このようにして計測したデータグローブ4の出力と計測に用いたロボットハンドの時系列関節角度パターンとを重回帰分析にかけ、各関節ごとの偏回帰係数を導き出す。下記の式は本実施の形態で用いる重回帰式である。

[数1]

$$\theta_i(t) = a_{0n} + \sum a_{in} x_{in}(t)$$

- [0037] ここで、目的変数 θ_i はロボットハンド1での制御対象となる関節各度の指令値、説明変数 x_{in} はデータグローブ4からの出力値、 a_{in} は偏回帰係数、 a_{0n} は残差である。この重回帰式で出される値はデータグローブ4での各指の関節角度の値であるので、この値をロボットハンドの角度指令値として制御を行えばよいことになる。

- [0038] 図6には、事前収集動作指令を得る際に用いる多指ロボットハンドの制御装置と本発明のロボットの駆動方法を用いてロボットハンドを駆動する装置の構成の一例の概略構成を示してある。図6に示した動作指令を事前に収集する際に用いる制御装置では、関節角度を指示する関節角度指令値 y に応じて動作するアクチュエータを備

えた k 個(k は正の整数)の関節を有する多指ロボットハンド1を制御するものとする。そしてデータグローブ4は、操作者の手に装着されて操作者の手の動きに応じて n 個(n は正の整数)の出力信号 x_n を出力する遠隔指令装置を構成している。この多指ロボットハンドの制御装置は、対応関係決定手段6と制御信号発生手段7とを備えている。対応関係決定手段6は、 k 個の関節のための k 個の関節角度指令値 y それぞれについて、各関節角度指令値 y と n 個の出力信号 x_n との間の対応関係を予め定めるように構成されている。具体的には、対応関係決定手段6は、時系列関節角度パターン記憶手段61と、サンプリング手段62と、偏回帰係数導出手段63と、重回帰式決定手段64とを備えている。

[0039] サンプリング手段63は、 k 個の関節(実際にはアクチュエータ2)にそれぞれ予め定めた時系列関節角度パターン記憶手段61に記憶した時系列関節角度パターンに従って関節角度指令値を入力して多指ロボットハンド1を動作させた状態において、多指ロボットハンド1の動きに倣うように操作者がデータグローブ4を操作したときに n 個の出力信号 x_n を時系列で m 回(m は n より大きい正の整数)サンプリングする。サンプリング手段62はサンプリングを開始する前に時系列関節角度パターン記憶手段61に指令を与え、時系列関節角度パターン記憶手段61は駆動装置3に予め定めた関節角度指令値を出力する。なお時系列関節角度パターンは、 N 次元空間をまんべんなく動くような任意の関数指令を多指ロボットハンドの関節(アクチュエータ)に与えるように構成するのが好ましい。駆動手段3は、これを受けて多指ロボットハンド1を動作させる。

[0040] 偏回帰係数導出手段63は、時系列関節角度パターン記憶手段61に記憶した時系列関節角度パターンと、サンプリング手段62から出力される m 回のサンプリングにより求めた m 組の n 個の出力信号 x_n とに基づいて重回帰分析を行い、 k 個の関節のそれぞれについて偏回帰係数を導出する。

[0041] 重回帰式決定手段64、偏回帰係数導出手段63が定めた k 個の関節毎の偏回帰係数を用いて、 k 個の関節毎の関節角度指令値 y と n 個の出力信号 x_n との間の対応関係を定める k 個の重回帰式を決定する。

[0042] そして制御信号発生手段7は、対応関係決定手段6の重回帰式決定手段64が求

めた対応関係とn個の出力信号 x_n とに基づいて、k個の関節角度指令値 y を求め、k個の関節角度指令値 y を示す制御信号(動作指令)を多指ロボットハンドのk個の関節を駆動するための駆動装置3に出力するように構成されている。この制御信号発生手段7は、重回帰式決定手段64が求めたk個の重回帰式を対応関係として用い、n個の出力信号 x_n を入力としてk個の関節角度指令値 y を求める。これらk個の関節角度指令値 y が動作指令となる。なおデータグローブ4の動きに応じて制御指令発生手段7からは処理速度に応じた時間間隔で動作指令が出力される。なお図6の例においては、対応関係決定手段6及び制御信号発生手段7により動作物(操作者の手)に所定の動作を行わせたときの複数のセンサの出力に基づいて動作指令を発生する動作指令発生装置が構成されている。

- [0043] 重回帰式決定手段64で決定するデータグローブ4の出力と多指ロボットハンド1の関節角度との対応を示す重回帰式は以下のような形で記述することもできる。

[数2]

$$HAND_i = a_0[n] + \sum a_i[n] \cdot GLOVE_i[n]$$

- [0044] 上記式においてHANDは、多指ロボットハンド関節値、 a は偏回帰係数、 $GLOVE_i[n]$ はn自由度のデータグローブ出力、 i は多指ロボットハンドの関節数である。なお上記式を用いる場合には、行が列より大きくなり優決定となる。そのため特異値分解を用いて最適解としての偏回帰係数を求めることが好ましい。

- [0045] 図7は、図6の対応関係決定手段6の主要部をコンピュータを用いて実現する場合に用いるソフトウェアのアルゴリズムを示すフローチャートである。まずステップST1では、データグローブ4と時系列関節角度パターンのデータ対を獲得する。即ち多指ロボットハンド1の各関節可動域をまんべんなく動かすような関数を多指ロボットハンド1に与えて駆動し、多指ロボットハンド1に合わせて操作者の手指も同じ動作を行う。多指ロボットハンド1のある関節角度または関節角度指令値を y とし、n個のデータグローブ出力を X_n とすると、次式のように時刻1〜mまでの時系列データが得られる。

[数3]

$$\begin{aligned}
 y_1 &= x_1 a_{11} + x_2 a_{12} + \cdots + x_n a_{1n} \\
 y_2 &= x_1 a_{21} + x_2 a_{22} + \cdots + x_n a_{2n} \\
 &\vdots \\
 y_m &= x_1 a_{m1} + x_2 a_{m2} + \cdots + x_n a_{mn}
 \end{aligned}$$

[0046] ただし上記式において、 $m > n$ である。また前提として、 x にはシステム雑音、量子化誤差などの雑音が含まれるものとする。

[0047] 次にステップST2進んで、データ対をベクトルと行列の形式で表現すると、 $y = [H]$ x のように表現することができる。ただしこの式における y 、 x 及び $[H]$ は下記の通りである。

[数4]

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad [H] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & & \cdots & \\ \vdots & & & \vdots \\ a_{m1} & & \cdots & a_{mn} \end{bmatrix}$$

[0048] 次にステップST3で、最小二乗法による解 x^- の推定を開始する。この推定の過程において、偏回帰係数が求められる。最小二乗法に用いる式は以下の式である。

[数5]

$$x^- = ([H]^t [H])^{-1} [H]^t y$$

[0049] 次にステップST4から、ヤコビ法による $([H]^t [H])$ の計算を開始する。すなわち直交行列 P を使って $([H]^t [H]) (= [A])$ を簡単な形の類似行列 $[B]$ に変換する。言い換えれば、すなわち、 $[P]^{-1} [A] [P] = [B]$ を計算する。

[0050] 次にステップST5で、行列 $[A]$ の非対角要素のうちの絶対値最大のもの $a_{kl} (= a_{lk})$

を選ぶ。そして次にステップST6において、 $a_{kk} = a_{ll} = \cos \theta$ 、 $a_{kl} = -\sin \theta$ 、 $a_{lk} = \sin \theta$ とする単位行列[P]を作る。

[0051] 次にステップST7へと進んでcと θ の値を決定する。 θ は、下記の行列により求める。

[数6]

$$\theta = \left\{ \begin{array}{ll} 1/2 \arctan (2a_{kl} / (a_{kk} - a_{ll})) & a_{kk} \neq a_{ll} \\ \text{sign}(a_{kl}) \pi / 4 & a_{kk} = a_{ll} \end{array} \right\}$$

[0052] ただし、上記式において $k < l$ である。

[0053] 次にステップST8へと進んで、 $[P]^t [A] [P]$ を計算して新しい[A]とする。次にステップST9で、行列[P]の積から固有ベクトルVを下記の式で求める。

[数7]

$$V = [P]_{(m)}^t [P]_{(m-1)}^t \cdots [P]_{(2)}^t [P]_{(1)}^t$$

[0054] ただし、上記式において $[P]_{(N)}^t$ はN回目の変換に用いた行列[P]である。

[0055] 次にステップST10へと進み、[A]のどの非対角要素も収束条件 < 0.0001 なら終了し、そうでなければステップST5へと戻る。ここで行列[B]の対角要素が固有値となる。

[0056] 次にステップST11へと進んで解に含まれる雑音を抑制する。ステップST12では、固有値 $(\lambda_1, \lambda_2, \cdots \lambda_n)$ を $(1/(\lambda_1 + v), 1/(\lambda_2 + v), \cdots 1/(\lambda_n + v))$ と置き換えて逆行列 $([H]^t [H])^{-1}$ を求める。ただし $([H]^t [H])^{-1} = [U] [A]^{-1} [U]^t$ である。またベクトル u_i は $[H]^t [H]$ の固有ベクトルである。また $[U] = [u_1, u_2, \cdots u_N]$ である。更に $[A]^{-1}$ は下記のように表される。

[数8]

$$[A]^{-1} = \begin{bmatrix} 1/(\lambda_1 + v) & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1/(\lambda_2 + v) & & & & \\ 0 & & & & & \\ \vdots & & & & & \\ 0 & & & & & \\ & & & \cdots & & 1/(\lambda_N + v) \end{bmatrix}$$

[0057] ただし $v = \sigma_n^2 / \sigma_x^2$ である。ここで、雑音成分 n^- を $y = [H]x$ の右辺に加え、 $n^- = y - [H]x$ と変形することで σ_n^2 を計算することができる。また σ_x^2 は y から直接、計算が可能である。そしてステップST13で最小二乗法による解 x^- の推定を終了する。

[0058] 上記フローチャートのアルゴリズムに従えば、偏回帰係数を導出するために最小二乗法と特異値分解とを用いて重回帰分析を行うことになる。なお上記ステップST4ーST10までが特異値分解を実施しているステップである。

[0059] 図6と図8乃至図10を用いて、本発明の方法によりロボットハンドを駆動する実施の形態の一例の説明をする。本発明の実施の形態の第1のステップでは、事前収集動作指令記憶装置8が、データグローブ4を嵌めた手(動作物)に所定の動作を行わせたとときのデータグローブ4の複数のセンサ5の出力に基づいて動作指令発生装置(6, 7)が発生する動作指令を事前収集動作指令として記憶する。この記憶動作をコンピュータを用いて実現する場合に用いるソフトウェアのアルゴリズムを図8に示す。

[0060] 第2のステップでは、図9に示すように、単眼カメラまたは両眼カメラ等からなるカメラ9によりデータグローブ4を手に着用して事前収集動作指令を取得するときと同じ動作(所定の動作)を人の手Hで行う。そしてこのときの手(動作物)Hの画像データを時系列で取得する。なおこの画像データの取得は、事前収集動作指令を取得するのと同様に行ってもよいが、後から行ってもよい。図9の例では、手Hで「グー」の形を作った状態の画像データを取り、この画像データに基づいてコンピュータグラフィック技術で手の擬似物の画像データPD1を作成する。そして手Hで「パー」の形を作った状態の画像データを取り、この画像データに基づいてコンピュータグラフィック技術で手の擬似物の画像データPDnを作成する。「グー」から「パー」まで変化するまでの

間の過程の画像を実際に撮影してもよいが、図9のようにその間の画像をコンピュータグラフィック技術で作成してもよい。このようにして作成した画像データは、画像データ記憶装置10に記憶される。

[0061] 第3のステップでは、第2のステップで得た複数の画像データに含まれる各画像データと事前収集動作指令記憶装置8に記憶された事前収集動作指令とを対応付けて画像対応動作指令記憶手段11に記憶する。なお本例では、画像対応動作指令記憶手段11は、実際に撮影した画像だけでなく、時系列で取得した前の画像データPD1と後の画像データPDnとの間の未作成の画像データをコンピュータグラフィック技術により作成する場合には、作成した画像データに対応する事前収集動作指令を前の画像データPD1に対応する事前収集動作指令と後の画像データPDnに対応する事前収集動作指令とに基づいて推定により作成して未作成の画像データと対応付けて記憶することができる。このようにすれば、少ない画像データによりより多くの画像データと事前収集動作指令との対応データを得ることができ、基礎データ(画像データと事前収集動作指令とを対応付けたデータ)の収集が大幅に簡単になる。

[0062] 第1から第3までのステップは、ロボットハンドを駆動するための事前作業として実施される。なお自由度ごとに5度刻みの精度の関節角度制御をロボットハンドの駆動において実現しようとする、1億個(20の7乗)程度の数の画像が必要となるが、上述のようにコンピュータグラフィック技術を用いれば、基礎データの収集は簡単になる。この場合に用いることができるCG編集ソフトウェアとしては「Poser 5(Curious Labs Incorporated製)」を使用することができる。このソフトウェアを用いれば、データグローブ出力から得られた関節データから、撮像した手指2次元画像と類似した理想的な手指CG画像を補間生成することができる。そして補間生成した画像データに対応する事前収集動作指令は、関節角度の線形補間により生成することができる。なお、画像1枚あたりの手指CGのファイルサイズは5kB程度である。

[0063] 第4のステップ及び第5のステップは、実際にロボットハンドを駆動する際のステップである。第4のステップでは、ロボットハンドを動作させるために、手で所望の動作を行い。図6に示すように、その際の手の画像データをカメラ12を用いて時系列でロボット動作用画像データとして取得する。このロボット動作用画像データ(動作画像デ

ータ)は、ロボット動作画像データ記憶装置13に記憶される。そして第5のステップでは、ロボット動作画像データに含まれる動作画像データに対応する画像データを画像対応動作指令記憶手段11に記憶している複数の画像データから時系列で特定する。この特定作業は、画像データ特定及び動作指令発生装置14で行う。この例では、ロボット動作画像データに含まれる動作画像データと画像対応動作指令記憶手段に記憶されている複数の画像データとの対応は、両者の類似度に基づいて判断する。類似度の求め方は任意である。そして第5のステップでは、特定した画像データに対応する事前収集動作指令を動作指令として駆動装置3に与え、駆動装置3の出力で多指ロボットハンドを駆動する。第4及び第5のステップの主要部をコンピュータ用いて実現する場合のソフトウェアのアルゴリズムは図10に示す通りである。

- [0064] 画像データ特定及び動作指令発生装置14で行う画像データの特定の際に特徴量を抽出する技術として、例えば、高次局所自己相関関数を用いる特徴量抽出技術を利用することができる。高次局所自己相関関数については、1993年発行の電総研究報告、j957の126頁乃至140に「柔らかな情報処理のための統計的手法の応用に関する研究」と題する論文に詳しく説明されている。手指2次元画像の特徴抽出には、高次局所自己相関関数を用いた。画面内の対象画像を $f(r)$ とすると、 N 次自己相関関数は変位方向(a_1, a_2, \dots, a_N)に対して次式のように定義される。

[数9]

$$x^N(a_1, a_2, \dots, a_N) = \int f(r) f(r + a_1) \cdots f(r + a_N) dr$$

- [0065] 実証試験では、高次自己相関係数の次数 N を2とし、また変位方向を参照点 r の周りの局所的な 3×3 画素の領域に限定した。これより、平行移動による等価な特徴を除くと特徴 M の数は図11に示す25個になる。図11中の黒の四角は、局所パターンの対応する画素の位置に対応する。各特徴の計算は、局所パターンの対応する画素の値の積を全画素に対して足し合わせることで得られる。

- [0066] 上記の方法によれば、実際にロボットハンド1を動かすために、ロボットハンド1に対

応する手の画像を得るだけで、センサを使用することなく、手の動きと同じ動作を、見まねにより不良設定性を解消してロボットハンドに動作させることができる。

[0067] 上記実施の形態では、第3のステップで実際に人の手の画像を撮影しているが、すべてコンピュータグラフィック技術により手の画像データ(手の擬似物の画像データ)を作成してもよいのは勿論である。

[0068] 次に上記実施の形態を改善または改良する手法について詳しく説明する。

[0069] [個人差の考慮]

まず前述の第2のステップにおいて、複数の画像データ(コンピュータグラフィックの画像データに、人間の手に現れる個人差を考慮した画像データを含めておくと、第5のステップにおける画像データの特定精度を高めることができる。図12(A)乃至(C)は、人の手に現れる個人差の相違要素を示している。図12(A)は手をグーの形に握った状態を示している。拇指以外の4指の基節位置の曲率に個人差が現れる。すなわち4指の基節位置が直線的になる手もあれば、湾曲した形状になる手もある。また拇指基節の張り出し具合にも顕著な個人差が現れる。これらの個人差は、2つの関節間の長さの相違によって生じる。また図12(B)は手をチョキの形にした場合の状態を示している。この状態においては、二本の指の開き具合と開いた二本の指の基準角度の違いに個人差が現れる。更に図12(C)は手をパーの形に開いた状態を示している。この状態では、拇指と人差し指との間の開き具合と拇指の反り具合に個人差が大きく現れる。このような個人差は、コンピュータグラフィック技術を用いて手の形の画像データ(疑似画像データ)を作成する場合に、相違要素をパラメータとして既に作成した基本となる画像データを変更することにより簡単に得ることができる。図13は、コンピュータグラフィック編集ソフトのパラメータを変えて、基節位置の曲率と拇指基節の張り出し具合を変えた場合の画像データの相違を示す図である。この2つの相違要素を変更するだけでも、かなり異なった個人差の画像データを得ることができる。なおこれらの画像データを作成する際に、新たに作成する未作成の画像データに対応する事前収集動作指令を、前の画像データに対応する事前収集動作指令と後の画像データに対応する事前収集動作指令とに基づいて推定により作成し、これを未作成の画像データと対応付けて記憶することができるのは勿論である。

[0070] [解像度の変更]

次に第2のステップで作成する複数の画像データには、複数の画像データの解像度を変えて作成した複数の解像度変更画像データを含めてもよい。これは画像データの解像度を変えたほうが、画像の特徴が現れやすくなるものもあるからである。画像データの解像度は、基本となる画像データがあれば、簡単に変更することは可能である。図14は解像度を変えた3種類の画像データを示している。左端の画像は、原画像であり、中央の画像は、縦2画素×横2画素を1画素にして解像度を下げた状態の画像であり、右端の画像は、縦4画素×横4画素を1画素として更に解像度を下げた画像である。例えば、原画像が15,000枚あるとすると、解像度を2段階変更することにより、45,000種類の画像データが得られることになる。手をグーの形に握っている状態では、指が重なり合うため、解像度が高いほうが特徴量を正確に把握することができる。しかし手をチョキの形にしている状態では、立っている二本の指の部分から得られる画像の特徴量が重要であるが、その他の部分の特徴量がかえって誤った判断をする原因となる。そこで解像度の異なる画像データを予め基礎データとして用意しておけば、照合の精度は高くなる。

[0071] [各画像データにおける特徴量の計算]

先に、画像データの特定のために、高次局所自己相関関数を用いる特徴量抽出技術を用いる点について説明した。従来の技術では、画像全体について特徴量を抽出しているが、精度を高めるために画像をいくつかの分割画像に分けて分割画像単位で特徴量を抽出することが好ましい。これを実現するためには、まず画像データにおける手の重心を推定する。そのためには、コンピュータグラフィックの画像データを二値化して、白黒画像とする。そしてこの白黒画像について、下記式を用いてX軸方向、Y軸方向のそれぞれで中心を求める。

[数10]

$$x_g = \frac{1}{k} \sum_{i=1}^{160} x_{i*2}$$

$$y_g = \frac{1}{k} \sum_{i=1}^{120} y_{i*2}$$

[0072] 上記式においては、 x_g 、 y_g は重心のX、Y座標であり、 x_i 、 y_i は白画素のX、Y座標であり、 k は白画素の数である。

[0073] このようにして画像の中心を求めた後は、この重心をもとに各画像を画面分割する。図15(A)及び(B)は、 320×240 画素から構成された1画面を 4×4 画面に分割した状態と、 8×8 画面に分割した状態を示している。なお上記画素数であれば、 16×16 画面に分割することも可能である。

[0074] このようにして分割した各分割画像について、特徴量の計算を行う。即ち画面分割ごとの高次局所自己相関パターンを算出する。先に説明した高次局所自己相関パターンの求め方に従って画面分割ごとの高次局所自己相関パターンを算出すればよい。下記の式は高次局所自己相関パターンの算出に用いる式である。

[数11]

$$x^N(a_1, a_2, \dots, a_N) = \int f(r) f(r + a_1) \cdots f(r + a_N) dr$$

[0075] 上記式において、 x^N はN次元の参照点 r の周りの局所的な相関関数である。したがってNは2である。 (a_1, a_2, \dots, a_N) は、N次自己相関関数の変位方向であり、 $f(r)$ は画面内の対象画像の画素位置 r の輝度値である。図12に示した局所パターンを用いれば25パターンとなるが、重複を認めれば35パターンを使用してもよい。

[0076] [特徴量低減]

上記のように特徴量を計算した場合、画像データが多くなると処理すべき特徴量が

多く

なりすぎる。例えば、図14に示したように3つの解像度の画像データを利用する場合の特徴量は、特徴量＝解像度3×画面分割64×高次局所自己相関パターン25＝4800(次元)となる。そこで公知の主成分分析技術を利用して特徴量を低減することが好ましい。まず主成分分析による各主成分ごとの因子負荷量を算出する必要がある。下記の式は、主成分分析で用いる主成分得点の演算式である。

[数12]

$$Z_{kp} = \sum_{l=1}^{div} \sum_{m=1}^{pnum} \sum_{n=1}^{25} a_{klmn} x_{plmn}$$

[0077] 上記式において、

Z_{kp} は第k主成分におけるデータpの主成分得点であり、

x_{plmn} はデータpの第l解像度(lはアルファベットのLの小文字)の第m画面の第n特徴量であり、

a_{klmn} は第k主成分の第l解像度の第m画面の第n特徴量の因子負荷量であり、

divは解像度の数(たとえば3)で有り、pnumは画面分割数(たとえば8×8)である。

[0078] ここで因子負荷量 a_{klmn} を算出して上記式に代入すれば、第k主成分におけるデータpの主成分得点を求めることができる。なお因子負荷量 a_{klmn} の求め方には、主因子法、最小2乗法、最尤法などがある。主因子法で因子負荷量を求める場合には、次のようににする。

[0079] (1) 各変量の分散を1に標準化して、標本相関行列 $C = (r_{jk})$ を求める。

[0080] (2) x_j と残りの変量との相関係数のなかの最大値 $r_{\max}(j)$ (ただし $j \neq k$) を、Cの対角要素に代入して C^* を作る。

[0081] (3) C^* の固有値問題を解き、固有値 $\lambda_1, \lambda_2, \dots, \lambda_p$ と、それに対応する固有ベクトル c_1, c_2, \dots, c_p を求める。この固有ベクトルが因子負荷量に相当する。

[0082] 次に特徴量低減のために、各主成分の寄与率を求める。ここで寄与率とは、各主成分が元の情報をどれだけ説明しているかを表す係数であり、下記のように表すことができる。

[数13]

$$C_k = \sum_{l=1}^{div} \sum_{m=1}^{pnum} \sum_{n=1}^{25} b_{klmn}^2 / (div * pnum * 25)$$

[0083] ここで、 C_k は第k主成分の寄与率であり、

b_{klmn} は主成分得点 Z_{kp} と x_{plmn} の相関係数で、次式のように定義される。

[数14]

$$b_{klmn} = a_{klmn} \sqrt{\lambda_k}$$

[0084] a_{klmn} は第k主成分の第l解像度の第m画面の第n特徴量の因子負荷量であり、
 λ_k はk番目に大きい相関行列の固有値である。

[0085] 次に累積寄与率に基づいた主成分の数の決定を行う。前述の寄与率 C_k は、各主成分が元の情報をどれだけ説明しているかを表す係数であり、以下の関係式が成り立つ。

[数15]

$$C_1 \geq C_2 \geq C_3 \geq \dots \geq C_k \geq C_{k+1} \geq \dots \geq C_{div * pnum * 25}$$

[0086] そこで、累積寄与率95%程度を目処に、特徴量低減で使用する主成分の数を下記の式で決定する。

[数16]

$$\sum_{k=1}^{div * pnum * 25} C_k \geq 0.95$$

[0087] 本実施の形態では、累積寄与率が約97%となる第10主成分までを用いることにする。

[0088] 次に主成分得点に応じたデータの並び替えを行う。第1から第10主成分までの各

主成分ごとに、主成分得点の大きさに応じて全データを並べ替え(ソート)する。この並べ替えによって10種類の画像データソースが作成されることになる。図16は、並べ替えのイメージを示している。このような並べ替えによって、オリジナルの画像数が15,000枚とすれば、 $15,000 \times 10 = 150,000$ 枚の画像データが記憶されることになる。

- [0089] なお未知のヒト手指画像が入力された場合に、効率よく類似画像を検索するため、範囲を限定した検索対象数をあらかじめ決めておくのが好ましい。具体的には、各主成分において未知画像と同じか、最も近い主成分得点 Z_{kp} を持つデータと、各主成分の寄与率に応じた数の前後データを検索対象とする。各主成分の候補数は、以下の通りである。

[数17]

$$dc_p = Cc * \frac{\lambda_p}{\sum_{i=1}^{10} \lambda_i}$$

- [0090] ここで、 dc_p は第 p 主成分($p=1, 2, \dots, 10$)の推定候補数である。
- [0091] Cc は推定候補数の総和で、あらかじめ決めておくものである。
- [0092] そして λ_p を含む分数式は、第10主成分中の p 成分の寄与率である。
- [0093] 上記の式によって各主成分の寄与率に応じた数の前後データが検索対象となることが決まる。例えば、図16を参照して説明すると、照合対象となる動作画像データの第1主成分の主成分得点 Z_{kp} が13点であるとすると、第1主成分の主成分得点の得点順に並べ替えた基礎となる画像データのうち主成分得点が13点に近い画像データを中心に含んで、前述の推定候補数の画像データが照合用画像データとして抽出される。
- [0094] 上記の改良点を含めた画像データのデータベース作成の手順(実施の形態の第1のステップから第3のステップに相当)を図17に示す。
- [0095] 次に上記のようにして作成したデータベースを利用して、本発明の第4及び第5の

ステップ(照合処理)を主成分分析技術を用いて処理する処理手順について説明する。図18は、具体的な照合手順の流れを示している。まず手の実画像の撮像を得る。前述の特徴量の低減技術を用いると、ビデオレート以上の速さの処理速度が出せるので、撮像には高速度カメラを使用することができる。不必要な情報を除去するために、撮像からは背景を除去する。そのためには(手映像+背景映像)の撮像から(背景映像)を差し引く処理を行う。

- [0096] 次に撮影した画像(動作画像データ)の特徴量を計算する。撮影したときの解像度そのまま特徴量を計算してもよいが、図18の例では、撮像した動作画像データの解像度を変えて3種類(複数種類)の解像度の異なる動作画像データを得る。すなわち単純に見ると同じ動作画像データについて3倍のデータを作成する。そしてこれら解像度を変えた3種類の動作画像データ(手画像のデータ)について、前述のデータベースの作成時と同様に、手の画像の重心推定を行い、画面を分割する。この場合も、画面を16分割、64分割または246分割する。分割数は任意である。しかにあまり分割数を多くしても精度が上がらないことが確認されているので、16分割または64分割程度の分割数が好ましい。
- [0097] 次に各画面分割ごとの高次局所自己相関パターンを算出する。このときの高次局所自己相関パターンの算出も前述のデータベースを作成する際に用いた高次局所自己相関パターンの算出手法と同じである。
- [0098] 次に算出した高次局所自己相関パターンに基づいて主成分分析により主成分得点の計算を行う。この主成分得点の計算も前述のデータベース作成の際の主成分得点の計算方法と同じ計算方法を採用する。
- [0099] 次にデータベースに記憶されている基礎となる複数の画像データとの照合を行う。照合の候補となる照合用の複数の画像データを、計算により求めた主成分得点を基準にしてデータベースから選出する。なお各主成分からの候補数は、前述の[数17]の式により、あらかじめ決められている。そこでそれぞれの主成分において最も近い主成分得点 Z_{kp} を持つ画像データと、この画像データの前後の画像データを各主成分の寄与率に応じて予め定めた候補数の数分選出して照合用の画像データとする。データベースに記憶されている10種類の主成分を基準にして並べ替えられた10種

類の画像データソースから、それぞれ予め定められた候補数の画像データが選出される。仮に各主成分について予め候補数が20と定められているものとする、 $20 \times 10 = 200$ 枚の画像データがデータベースから選出されることになる。これらの画像データ中には重複しているものがあるのは勿論である。

- [0100] 次に類似度の計算を行う。具体的には、入力画像(動作画像データ)と候補となった複数枚の照合用の画像データ(候補CGハンド画像)との間で、次式により、ユークリッド距離 E_r の計算を行う。

[数18]

$$E_r = \sum_{i=1}^{10} (f_i(x_r) - f_i(x_t))^2$$

- [0101] ここで、 $f_i(x)$ は特徴量から算出した第 i 主成分の主成分得点であり、
 x_r は候補 r の高次局所自己相関関数による特徴量であり、
 x_t は時刻 t における高次局所自己相関関数による特徴量であり、
 E_r を最小にするデータ p を検索画像とし、データ p が持つ関節角度データを推定角度とする。

- [0102] 時刻 t と時刻 $t-1$ の検索結果が許容される関節角度の範囲であれば、時刻 t での探索を終了する。そして時刻 $t-1$ から大幅に異なる関節角度が選ばれている場合は、ユークリッド距離 E_r が次に小さい次候補を選び、許容範囲かどうかを下記の式で計算する。

[数19]

$$A_p = \sum_{i=1}^{24} (ang_{i(t)} - ang_{i(t-1)})^2$$

- [0103] ここで、 A_p は許容値であり、 i はデータグロブの関節番号である(本例では $i=24$)。

- [0104] また $ang_{i(t)}$ は時刻 t における i 番目の関節角度データである。

- [0105] 許容値が許容は範囲内であれば、その画像データが動作画像データに対応するものと特定される。
- [0106] 次に、本発明の方法の別の実施の形態について説明する。この実施の形態では、前述の実施の形態とは異なって、図19に示すように、データグローブを嵌めた手の上にさらに無地(無模様:一色)の手袋(被覆物)を嵌めて、前述の第1のステップと第2のステップとを同時に実施する。図19において、下側の図がデータグローブを嵌めた手であり、上側の図がデータグローブを嵌めた手の上に手袋を嵌めた図である。図20は、事前にデータベースを作成する際の手順を示すフローチャートであり、図21は図20のフローチャートの詳細を示すフローチャートである。
- [0107] まずビデオカメラ109で、手袋を嵌めた状態の手を撮影する。本実施の形態では、高速ビデオカメラ109で撮影した手画像データおよびデータグローブから得た角度データを用いてデータベースを構築する。そのために画像データと角度データとを組にして取得する。本実施の形態では、モノクロ高速度カメラ(MEGAPLUS社製、ES310/T)を使い、解像度を320*240画素として、手指が画面内に十分な大きさで映っている状態を想定する。また角度データの測定には、Virtual Technologies社製の「CyberGlove」(商標)と呼ばれるデータグローブを用いた。そして手袋は、白一色で無模様のものを用いた。手袋をデータグローブの上に嵌めることにより、データグローブ固有の画像特徴(センサの形状やグローブの形状)が出るのを防止している。前の実施の形態と異なって、白手袋を撮影すると、ヒトの手とは色が違い、画像としても違うものとなる。しかしながら、手袋を嵌めた手の画像データであっても、後に説明するように、特徴量化により一般の手画像として扱うことができる。なおデータベース作成に際しては、背景画像が影響を及ぼさないようにするために、背面に無模様一色のスクリーンを用いて撮影を行う。そして連続する複数の画像データと一緒に、データグローブから取得した手指角度データを組にして1次データベースを作成する。必要十分な情報を得るために、手を連続的に動かして必要な画像データと角度データを同時に取得して、保存する。
- [0108] 本実施の形態では、1次データベースの画像データを、実際に撮影した手画像の推定に必要な特徴量へと変換する画像処理を行なう。この画像処理のために背景除

去、解像度変更、エッジ抽出、輪郭抽出(特徴量抽出)を行なう。

[0109] まず画像情報から背景除去を行う。あらかじめ取得しておいた背景画像を用いて対象画像との差分を取る。その差分が閾値以上の場合は対象画像の画素を使い、以下ならば0とする。

[0110] 次に背景を除去した画像の解像度を変更する。今回取得した画像は320×240の画素値を持つ画像である。しかしこれをそのまま用いて各種画像処理を施して推定処理をするには、計算に時間がかかる。また人間が推定する際には、本実施の形態で用いた画像のサイズより小さくとも推定が可能である。そこで本実施の形態では、解像度を、320×240の画素値から64×64のサイズに変更する。解像度変更後の画素値は次式により得られる。

[数20]

$$gr(i, j) = \frac{1}{r} \sum_k \sum_l go(i * 320 / 64 + k, j * 320 / 64 + l)$$

[0111] ここでgr(i, j)は解像度変更後のi行j列番目の画素値である。go(i, j)は、解像度変更前のi行j列番目の画素値である。この式では、縦についても320で計算しているが、これは変更後の画素値を64×64にしているため縦横比を合わせるためである。またk, lについてはi×320／64から(i+1)×320／64-1まで計算し、rはkの個数×lの個数とする。

[0112] 以上の操作で得られた解像度変更画像を用いてエッジ抽出を行なう。この作業で推定に必要な特徴を抜き出す。具体的にはソーベルフィルタで解像度変更画像をフィルタ処理する。その後、この画像の最外周のみを二値化することでエッジを抽出する。

[0113] ソーベルフィルタによる処理は、次式で表現できる。

[数21]

$$fs(i, j) = \begin{cases} 127 & (\text{ただし } fr(i, j) > 70) \\ 0 & (fr(i, j) \leq 70) \end{cases}$$

[0114] ここで $fr(i, j)$ は下記の式で表すことができる。

[数22]

$$fr(i, j) = \sqrt{\frac{\{-2 * fi(i, j-1) - fi(i+1, j-1) - fi(i-1, j-1) + 2 * fi(i, j+1) + fi(i+1, j+1) + fi(i-1, j+1)\}^2}{+ \{-2 * fi(i-1, j) - fi(i-1, j+1) - fi(i-1, j-1) + 2 * fi(i+1, j) + fi(i+1, j+1) + fi(i+1, j-1)\}^2}}$$

[0115] 上記式において、 $fi(i, j)$ は処理前の横 i 番目、縦 j 番目の画素を示す。 fs が処理後の最終的な値となる。

[0116] ここでの画像の最外周の二値化は、輪郭抽出で輪郭が途切れないように、比較的低い閾値で二値化することを意味する。この最後の画像処理として、エッジ抽出で得られた二値化画像上を走査することで輪郭抽出を行う。前段階のエッジ抽出により得た画像を輪郭として使用しない理由として、環境光の変化を上げることができる。実際の推定時とデータベース作成時とでは、環境光や人の肌の色等に違いがある。ましてやデータベース作成時には白手袋を使用しているため、エッジ抽出を行っても、エッジの値や幅などが実際の画像のエッジとは異なってくる。その結果、データベースの画像データと実際の画像データとを比較するとき(推定時)とで、両者のエッジには大幅な相違が生じる可能性がある。そこで輪郭抽出を行って、エッジの幅や値を一定にする。輪郭抽出は、データベースから得られる情報と推定時に得られる情報との差を少なくし、環境の影響などを排除することで、推定時の誤差を減少させるために採用する。図22は、エッジ抽出後に輪郭抽出を行って得た解像度変更後の画像である。

[0117] なお最外周のみ二値化するための輪郭抽出アルゴリズムでは、画像を上から走査して行き、最初の127の点を始点とした。始点から左回りに近傍画素を探索して行き、127の点があればそこへ移動し、値を255に書き換える。移動後は前回の位置を参考にして探索開始地点を8パターンに分けて探索を継続する。始点に再び戻った時点で輪郭を抽出したと判断し、探索を終了させる。最後に残りの点は輪郭ではないので127の点は全て0とする。

[0118] 次に輪郭抽出を行った画像データの特徴量化を実施する。本実施の形態では、Otsu等(N. Otsu and T. Kurita, "A new scheme for practical, flexib

le and intelligent vision systems, "Proc. IAPR. Workshop on Computer Vision, pp. 431-435, 1998.)が提案している高次局所自己相関関数を特徴量化に使用した。高次局所自己相関特徴は、画像の認識や計測のために有効な特徴であり、下記の式で定義される高次相関関数について、参照点とその近傍についての計算をしたものである。

[数23]

$$x^N(a_1, a_2, \dots, a_N) = \int f(r)f(r+a_1)\cdots f(r+a_N)dr$$

[0119] 上記式で x^N はN次元でのr点近傍の相関関数である。一般に自然画像を処理対象とする場合には、対象点回りの画素が重要になるため、係数Nを2次までとする。また、平行移動すると等価なものになるものを除くと、図23に示す25の特徴量で表現できる。ただしNo. 1からNo. 5までは、No. 6以降の特徴量に比べスケールが小さいので、スケールを合わせる必要がある。No. 2〜No. 5までは、参照点の画素をさらにかけあわせ、No. 1は参照点の画素の2乗をかけることで他の特徴量にスケールを一致させる。

[0120] 本実施の形態では、図24に示すように、画像の画面を縦8分割、横8分割にした。そして、計64分割した分割画面に対して、それぞれ画素全てを参照点として、高次局所自己相関関数によって25の特徴量に次元低減を行なった。すなわち、全体の画像で考えれば、 $25 \times$ 画面分割数の特徴量により、その全体画像の特徴量化を行うことができる。

[0121] このようにして1つの画像データを特徴量化した画像データとデータグローブからの角度データ(事前収集動作指令に対応)を組みにした特徴データを画像対応動作指令記憶手段(データベース)に記憶する。本実施の形態では、このデータベースを利用して、実際に撮影した画像の特定を行う。

[0122] 次に、上記のようにして作成したデータベース(画像対応動作指令記憶手段)を利用して、実際に撮影した手の画像から手指角度を推定する方法について説明する。

図25は推定時の手順を示すアルゴリズムであり、図26はより詳細にアルゴリズムである。まず推定時には、データベース作成時と同様に高速カメラで動く手の画像(動作画像データ)を取得する。そして取得した動作画像データに対し、データベース作成時と同様の画像処理を施し特徴量を抽出する。そしてこの動作画像データの特徴量とデータベースに記憶されている複数の画像データの特徴量とを後述の処理により比較し、動作画像データと最も近い画像データと組になる角度データを、その動作画像データに対応する動作指令として出力する。

- [0123] 特に、「データベースとの照合」処理では、データベースに記憶されている画像データ全てに対して、動作画像データの特徴量との間の距離算出を行なった。距離の算出は単純なユークリッド距離で行なった。最終的な距離は、下記式より算出した。

[数24]

$$E_r = \sum_{i=1}^{25 \times n} (x_{ri} - x_{ti})^2$$

- [0124] 上記式において、 x_n は候補rの高次局所自己相関関数による特徴量i、 x_{ti} は時刻tでの高次局所自己相関関数による特徴量iである。Erを最小にするデータrの角度を、その動作画像データの推定角度とした。ただし、角度の大幅に違うものの中に特徴量の近いデータが紛れ込むことがあるので、前回(t-1時刻)の推定角度との比較を行って角度制限を行う。具体的には、距離が近いものについて、下記式で示す評価関数Apが規定値を超える場合には、推定候補から除外した。そして評価関数Apが規定値以内であれば、その推定角度を採用することとする。

[数25]

$$A_p = \sum_{i=1}^{24} (ang_{i(t)} - ang_{i(t-1)})^2$$

- [0125] ここで $ang_{i(t)}$ は時刻tのi番目の角度データである。また、データグロブから得られる角度情報は24となっている。以上の操作により現在の推定角度を求める。
- [0126] 実際の推定時に、データベースに記憶されている全ての画像データとの全てについて距離算出の処理を行なうと、データベースの大きさによっては計算コストがかかり

すぎ、実時間処理が不可能になる。そこで本実施の形態では、データのクラスタリングにより推定候補数の削減を行なうことにより、実時間処理を可能にする。

[0127] 図27は、クラスタリングを説明するための概念図である。図27においては、一つの行は複数枚の画像データによって構成される一つの組を示している。着色の濃い行が最初の探索対象となるデータ、着色の薄い行が詳細な探索対象となるデータである。データベースに含まれるデータを、先に説明したデータベース作成時のアルゴリズムで作成すると、画像データの並びにおいて、手の形状の近いものが集まるようになる。これを利用することで距離計算の対象画像データを削減する。具体的には複数枚の画像データを一組と考える。そして、始めの距離計算時に使う画像データは、複数組の中からそれぞれ1つを選び、選んだ画像データと距離計算を行う。これにより距離計算による計算コストは、各組に含まれる枚数-1がすべての組において削減されることになり、大幅なコスト減となる。ただし、この段階では真に近い角度の画像データも計算から除いている可能性があるので、推定結果が元のデータベースの画像データ全と対比する場合に比べ精度が落ちる。そのため1段階目の距離計算で距離の近いものについては、各画像データが含まれる組内のすべての画像データと動作画像データとの間の距離計算を行なう。このようにすれば各データ近傍には角度的に近いものが集まっているため、より精度の良い推定結果が得られる。クラスタリングの効果を確認する実験では、10枚の画像データを一組と考え、1枚の動作画像データについて1段階目の距離計算を複数組内の1枚の画像データとそれぞれ行なった。次にこの操作で得られた距離が近い上位の組5組に含まれる全ての画像データについて距離計算と角度制限を行ない、最終的な角度を推定した。その結果、すべての画像データと比較を行う場合と比べて、遜色の無い比較精度が得られることが確認できた。

[0128] 上記実施の形態では、特徴量を求めるにあたって、高次局所自己相関関数を用いた。しかしながらその他の特徴量の求め方を採用することができるのは勿論である。例えば、廣池敦、武者義則著の「大規模な画像集合のための表現モデル」(日本写真学会誌、66巻、1号、93-101頁、2003年発行)に記載の特徴量ベクトルの求め方を採用することもできる。

産業上の利用可能性

- [0129] 本発明によれば、実際にロボットを動かすために、ロボットに対応する動作物またはその擬似物の画像を得るだけで、センサを使用することなく、その動作物または擬似物と同じ動作を、不良設定性を解消してロボットに見まね動作させることができる。したがって本発明によれば、画像データを入力データとしてロボットを簡単に駆動することが可能になる。